Ebola outbreak in West Africa: real-time estimation and multiple-wave prediction

Xiang-Sheng Wang[a,*], Luoyi Zhong[a]

[a]*Department of Mathematics, Southeast Missouri State University, Cape Girardeau, MO 63701, USA*

Abstract

Based on the reported data until 18 March 2015 and numerical fitting via a simple formula of cumulative case number, we provide real-time estimation on basic reproduction number, inflection point, peak time and final outbreak size of ongoing Ebola outbreak in West Africa. From our simulation, we conclude that the first wave has passed its inflection point and predict that a second epidemic wave may appear in the near future.

Keywords: Ebola outbreak in West Africa, basic reproduction number, inflection point, final outbreak size, peak time, multiple epidemic waves
2000 MSC: 92B05, 62P10

1. Introduction

Ebola virus disease (EVD) is a severe disease in humans which has infected nearly 25 thousand individuals and claimed more than ten thousand deaths during the recent outbreak in West Africa, according to the report of World Health Organization dated 18 March 2015 [1, 2]. The most affected countries are Guinea, Liberia and Sierra Leone. This study aims to provide some real-time estimations on the outbreak in these three countries using the reported cumulative case data. Specifically, we will estimate the following quantities:

1. basic reproduction number R_0, which is defined as the average new cases caused by a single infective individual during one infectious period;
2. inflection point t_c, which marks the time when the increment speed of cumulative case numbers starts to slow down;
3. final outbreak size K, which indicates the total number of infectious cases throughout the outbreak wave.
4. peak time t_p, which is defined as the critical time when daily infectious number reaches its maximum.

All of these indicators provide quantitative information about severity of a disease outbreak.

2. Methods

Following [3], we study the epidemic model:

$$S'(t) = -\frac{\beta S(t)I(t)}{S(t) + I(t)};$$
$$I'(t) = \frac{\beta S(t)I(t)}{S(t) + I(t)} - \gamma I(t), \tag{1}$$

where $S(t)$ and $I(t)$ are the numbers of susceptible and infective individuals at time t, respectively. The constant β denotes the transmission rate of the disease, and the constant γ corresponds to the removal rate of infective individuals. The basic reproduction number [4, 5] is given by

$$R_0 = \frac{\beta}{\gamma}. \tag{2}$$

It is noted that a disease outbreak occurs if and only if $R_0 > 1$. The differential system (1) can be solved explicitly and its solution is given by

$$S(t) = K[1 + e^{\gamma(R_0-1)(t-t_c)}]^{-R_0/(R_0-1)};$$
$$I(t) = K[1 + e^{\gamma(R_0-1)(t-t_c)}]^{-1/(R_0-1)} \tag{3}$$
$$- K[1 + e^{\gamma(R_0-1)(t-t_c)}]^{-R_0/(R_0-1)},$$

---

*Corresponding author. Email: xswang@semo.edu

where K and t_c are two constants of integration. Now, we define the cumulative infective case number at time t as

$$C(t) = \int_{-\infty}^{t} \frac{\beta S(t)I(t)}{S(t) + I(t)}. \tag{4}$$

From (1) and (3), we have

$$C(t) = K - K[1 + e^{\gamma(R_0-1)(t-t_c)}]^{-R_0/(R_0-1)}. \tag{5}$$

Here, the constant $K = C(\infty)$ has the biological meaning of final outbreak size. It can be verified that $C''(t_c) = 0$. Hence, t_c is the inflection point of $C(t)$. We remark that the inflection point t_c is related to but different from another commonly used quantity: the peak time, denoted by t_p. The peak time is defined as the time when infective case number achieves its maximum, namely, $I'(t_p) = 0$. It follows from (3) that

$$t_p = t_c + \frac{\ln R_0}{\gamma(R_0 - 1)}. \tag{6}$$

In the case when R_0 is close to 1, namely, $\ln R_0 \approx R_0 - 1$, we can approximate the difference $t_p - t_c$ by $1/\gamma$. Thus, the peak time occurs about one infectious period after the inflection point [3].

Richards' empirical model [6] was suggested to provide real-time estimation of a disease outbreak; see [7] for example. However, some of the parameters in Richards' model do not have clear biological meanings [3]. The advantage of formula (5) is that all of the parameters in this formula have significant biological interpretations. We will use the explicit formula of $C(t)$ in (5) to fit the reported cumulative case numbers of 2014 Ebola outbreak in West Africa and provide real-time estimation of basic reproduction number R_0, inflection point t_c, final outbreak size K and peak time t_p.

As pointed out in [3], one should fix the value of γ, the removal rate of infective individuals, to resolve possible overfitting problems. Note that $1/\gamma$ can be regarded as the infectious period which characterizes the average duration of an individual being infective. In most cases, an individual is removed from the infective group either by recovery or death. For the fatal cases of Ebola virus disease, death usually occurs between 6 and 16 days (with mean 7.5 days) after onset of symptom; and for the non-fatal cases, patients may improve their symptoms at around day 6 but need more time to recover

[8]. Convalescent patients may still be infective because the Ebola virus RNA may remain in the body fluid for a couple of weeks even though the risk of transmission from them is low [9]. It is thus reasonable to assume the infectious period to be 7.5 days with some possible perturbations in the interval between 6 and 16 days. In our simulation, we first fix $1/\gamma = 7.5$ days to estimate the basic reproduction number R_0, inflection point t_c, peak time t_p and final outbreak size K using reported cumulative case data of Ebola virus in Guinea, Liberia and Sierra Leone, respectively [1, 2]. Also, we provide the 95% confidence intervals of each estimated parameter value using bootstrap method. Next, we vary the value of the parameter $1/\gamma$ from 6 to 16 days and investigate the sensitivity of fitted parameter values.

3. Results

The basic reproduction number R_0 is estimated as 1.116 (95% CI: 1.115-1.116) for Guinea, 1.226 (95% CI: 1.225-1.228) for Liberia, and 1.181 (95% CI: 1.181-1.182) for Sierra Leone. The inflection point t_c is estimated as 21 November 2014 for Guinea, 24 October 2014 for Liberia, and 28 November 2014 for Sierra Leone. As shown in Table 1, the lengths of 95% confidence intervals for the estimated inflection points are no more than one day.
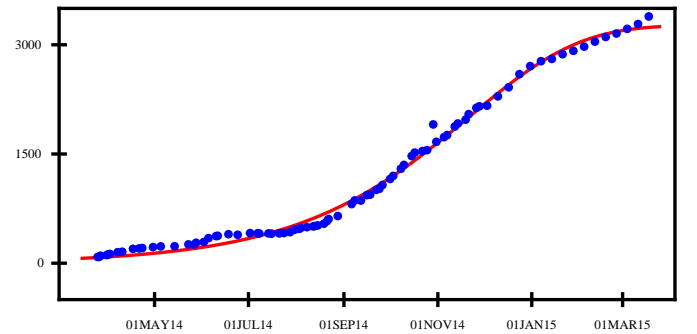

Figure 1: Fitted graph for the reported cumulative cases in Guinea. The dots are real data and the curve is plotted using fitted results.

The fitted curves together with reported cumulative case data are illustrated in Figure 1 (Guinea), Figure 2 (Liberia) and Figure 3 (Sierra Leone). It is noted that in each of these three figures, there is a jump on the reported cumulative case numbers in

	final outbreak size K	basic reproduction number R_0	inflection point t_c	peak time t_p
Guinea	3268 [3257, 3274]	1.116 [1.115, 1.116]	266 [265, 266]	273 [272, 273]
Liberia	8630 [8605, 8660]	1.226 [1.225, 1.228]	238 [237, 238]	244 [244, 245]
Sierra Leone	11227 [11198, 11253]	1.181 [1.181, 1.182]	273 [272, 273]	279 [279, 280]

Table 1: Estimated parameter values with 95% confidence intervals. Here, day 1 corresponds to 1 March 2014. So, days 266, 238 and 273 correspond to 21 November 2014, 24 October 2014, and 28 November 2014, respectively.
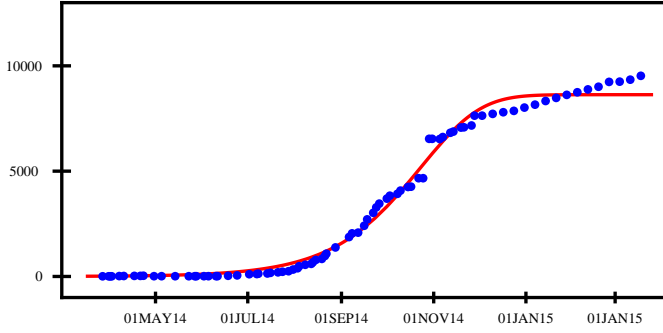

Figure 2: Fitted graph for the reported cumulative cases in Liberia. The dots are real data and the curve is plotted using fitted results.
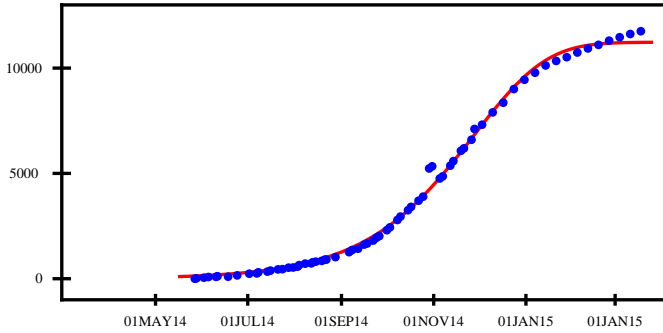

Figure 3: Fitted graph for the reported cumulative cases in Sierra Leone. The dots are real data and the curve is plotted using fitted results.

Now, we regularly increase the value of infectious period $1/\gamma$ from 6 to 16 days, and conduct numerical simulations. It is noted that the fitted values of basic reproduction number R_0 will also increase from 1.092 to 1.252 for Guinea, from 1.179 to 1.503 for Liberia, and from 1.144 to 1.400 for Sierra Leone; see Figure 4. The estimated final outbreak size stays in a range of [3261, 3307] for Guinea, [8620, 8677] for Liberia, and [11209, 11319] for Sierra Leone; see Figure 7. On the other hand, the inflection point t_c and peak time t_p do not vary too much; see Figures 5 and 6. For Guinea, t_c decreases from 266 (21 November 2014) to 264 (19 November 2014). For Liberia, t_c decreases from 238 (24 October 2014) to 235 (21 October 2014). For Sierra Leone, t_c decreases from 273 (28 November 2014) to 270 (25 November 2014). The peak time t_p increases from 272 (27 November 2014) to 279 (4 December 2014) for Guinea, from 244 (30 October 2014) to 248 (3 November 2014) for Liberia, and from 279 (4 December 2014) to 284 (9 December 2014) for Sierra Leone. We observe that the fitted inflection point t_c and peak time t_p are stable under perturbations on the infectious period $1/\gamma$.

late October 2014. This is due to a more comprehensive assessment of patient databases on the World Health Organization report dated 29 October 2014 [10]. Among these three countries, Liberia has the most significant gap, which may account for the result that the inflection point for Liberia is about one month earlier than the other two countries.

We also fit the final outbreak size as 3268 for Guinea, 8630 for Liberia, and 11227 for Sierra Leone. All of these estimated values are smaller than cumulative case numbers reported on 18 March 2015. This indicates that another potential outbreak wave may be approaching [7].
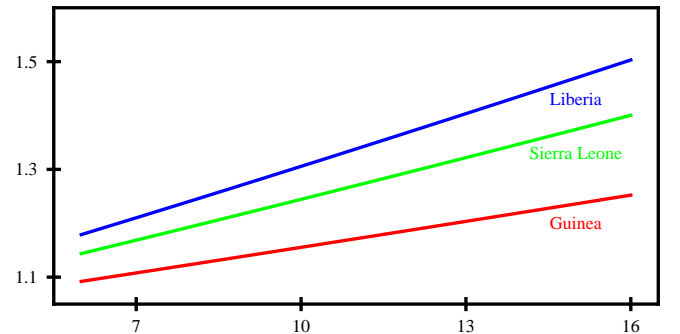

Figure 4: Estimated values of basic reproduction number when infectious period increases from 6 to 16 days.
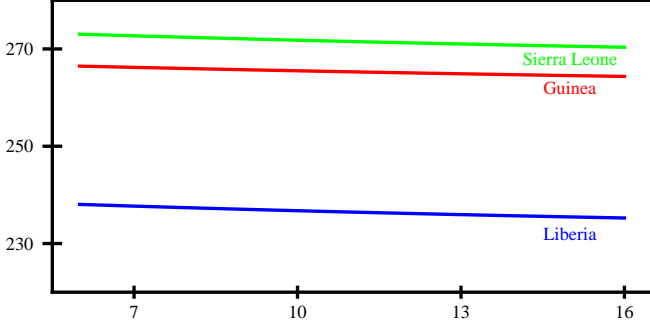
3

Figure 5: Estimated values of inflection point when infectious period increases from 6 to 16 days. Here, day 1 corresponds to 1 March 2014. So, days 266, 238 and 273 correspond to 21 November 2014, 24 October 2014, and 28 November 2014, respectively.
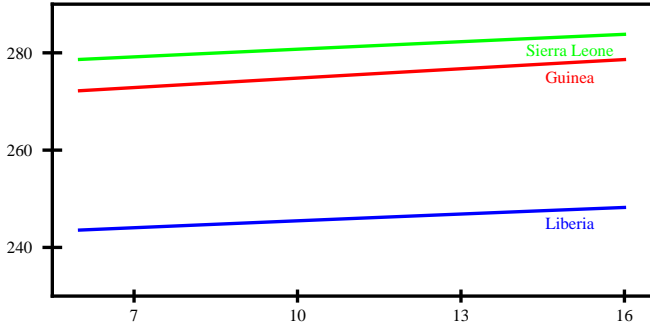

Figure 6: Estimated values of peak time when infectious period increases from 6 to 16 days. Here, day 1 corresponds to 1 March 2014. So, days 266, 238 and 273 correspond to 21 November 2014, 24 October 2014, and 28 November 2014, respectively.

4. Discussion

This study provides real-time estimation of basic reproduction number, final outbreak size, inflection point and peak time for the onging Ebola outbreak in West Africa using reported cumulative case data.

The fitted basic reproduction numbers are smaller than those estimated in [11] where only the data until 20 August 2014 was used. This indicates that the disease control policy became more effective during the late stage of the outbreak.

We also observe that the increment speed of cumulative case number began to slow down after 21 November 2014 in Guinea, 24 October 2014 in Liberia, and 28 November 2014 in Sierra Leone. The estimated inflection points for Guinea and Sierra Leone are close to each other, but the one for Liberia is about one month earlier. This is due to a signif-
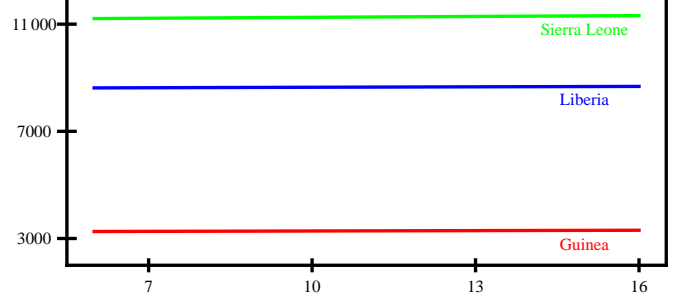

Figure 7: Estimated values of final outbreak size when infectious period increases from 6 to 16 days.

icant increase on reported cumulative case number dated 29 October 2014; see Figure 2 and [10].

From Table 1, we note that the estimated peak time has about one week's delay after the estimated inflection point, while the infectious period is fixed as 7.5 days. This supports the conclusion in [3] that the peak time occurs about one infectious period after the inflection point.

If we vary the infectious period from 6 to 16 days, the estimated basic reproduction number stays in a range of [1.092, 1.252] for Guinea, [1.179, 1.503] for Liberia, and [1.144, 1.400] for Sierra Leone. The estimated final outbreak size ranges from 3261 to 3307 for Guinea, from 8620 to 8677 for Liberia, and from 11209 to 11319 for Sierra Leone. The estimated inflection point and peak time are much stabler and only varies within a small interval. This demonstrates that our method has a significant accuracy in capturing the inflection point and peak time.

The values of final outbreak sizes in three countries are all underestimated, which can be considered as a warning signal of a second outbreak wave.

Appendix A. Explicit solution of (1)

Here, we provide the detail in solving the system (1). First, we add the two equations in (1) to obtain $S'(t) + I'(t) = -\gamma I(t)$. Coupling this with the first equation of (1) yields

$$\frac{d(S + I)}{dS} = \frac{\gamma(S + I)}{\beta S}.$$

This equation is separable and its solution can be written as

$$C_1(S + I) = S^{\gamma/\beta},$$

4

where $C_1 > 0$ is a constant of integration. Next, we use the above relation to eliminate I in the first equtaion of (1). It follows that

$$S'(t) = -\beta S(t) + \beta C_1 [S(t)]^{2-\gamma/\beta}.$$

This is a Bernoulli equation. We set $u(t) = [S(t)]^{\gamma/\beta-1}$ and obtain

$$u'(t) = (\beta - \gamma)[u(t) - C_1].$$

The above equation becomes separable and its solution is given by

$$u(t) = C_1[1 + C_2 e^{(\beta-\gamma)t}]$$

with $C_2 > 0$ being another constant of integration. Now, we have

$$S(t) = C_1^{\beta/(\gamma-\beta)}[1 + C_2 e^{(\beta-\gamma)t}]^{\beta/(\gamma-\beta)}.$$

Substituting this into the relation between S and I gives

$$\begin{aligned}
I(t) =& C_1^{\beta/(\gamma-\beta)}[1 + C_2 e^{(\beta-\gamma)t}]^{\gamma/(\gamma-\beta)} \\
& - C_1^{\beta/(\gamma-\beta)}[1 + C_2 e^{(\beta-\gamma)t}]^{\beta/(\gamma-\beta)}.
\end{aligned}$$

Finally, we set $K = C_1^{\beta/(\gamma-\beta)}$ and $t_c = -(\ln C_2)/(\beta - \gamma)$ to rewrite the formulas of $S(t)$ and $I(t)$ as

$$\begin{aligned}
S(t) =& K[1 + e^{(\beta-\gamma)(t-t_c)}]^{\beta/(\gamma-\beta)}; \\
I(t) =& K[1 + e^{(\beta-\gamma)(t-t_c)}]^{\gamma/(\gamma-\beta)} \\
& - K[1 + e^{(\beta-\gamma)(t-t_c)}]^{\beta/(\gamma-\beta)}.
\end{aligned}$$

This is equivalent with (3) in view of (2).

References

[1] World Health Organization. Situation reports with epidemiological data: archive. Available at `http://apps.who.int/ebola/en/current-situation`. (Last accessed on 18 March 2015.)

[2] Centers for Disease Control and Prevention. 2014 Ebola outbreak in West Africa - reported cases graphs. Available at `http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/cumulative-cases-graphs.html`. (Last accessed on 18 March 2015.)

[3] X.-S. Wang, J. Wu, Y. Yang, Richards model revisited: validation by and application to infection dynamics. Journal of Theoretical Biology. (313) 2012 12-19.

[4] O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, Journal of Mathematical Biology (28) 1990 365-382.

[5] P. van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Mathematical Biosciences (180) 2002 29-48.

[6] F. J. Richards, A flexible growth function for empirical use, Journal of Experimental Botany (10) 1959 290-300.

[7] Y. H. Hsieh, Y. S. Cheng, Real-time forecast of multiphase outbreak, Emerging Infectious Diseases (12) 2006 122-127.

[8] Centers for Disease Control and Prevention. Ebola virus disease (EVD) information for clinicians in U.S. healthcare settings. Available at `http://www.cdc.gov/vhf/ebola/healthcare-us/preparing/clinicians.html`. (Last accessed on 18 March 2015.)

[9] D. G. Bausch, J. S. Towner, S. F. Dowell, et al. Assessment of the risk of Ebola virus transmission from bodily fluids and fomites, The Journal of Infectious Diseases. (196) 2007 Suppl 2, S142-147.

[10] World Health Organization. Ebola response roadmap situation report - 29 October 2014. Available at `http://apps.who.int/iris/bitstream/10665/137376/1/roadmapsitrep_29Oct2014_eng.pdf?ua=1`. (Last accessed on 18 March 2015.)

[11] C. L. Althaus, Estimating the reproduction number of Ebola virus (EBOV) during the 2014 outbreak in West Africa, PLOS Currents Outbreaks. 2014 Sep 2. Edition 1. doi: 10.1371/currents.outbreaks.91afb5e0f279e7f29e7056095255b288.